

МАТЕМАТИЧЕСКИЕ И ИНСТРУМЕНТАЛЬНЫЕ МЕТОДЫ ЭКОНОМИКИ

УДК 338

*Е.Н. Барышева**

ИЗУЧЕНИЕ ЗАВИСИМОСТИ ПОКАЗАТЕЛЕЙ, ХАРАКТЕРИЗУЮЩИХ ДТП РАЙОНОВ САМАРСКОЙ ОБЛАСТИ, СРЕДСТВАМИ КЛАСТЕРНОГО АНАЛИЗА

В статье дается оценка дорожно-транспортных происшествий, совершению которых способствовали неудовлетворительные условия содержания и обустройства улично-дорожной сети в 33 районах Самарской области, методами кластерного анализа.

Ключевые слова: ДТП, Самарская область, регион, районы, кластерный анализ.

Дорожно-транспортная безопасность является одним из главных предметов озабоченности для многих институциональных общественных структур. К этим институциональным структурам относятся: правительство, система здравоохранения, полиция, СМИ, научное сообщество и другие. У различных пользователей информации о дорожно-транспортной безопасности существуют разные интересы и потребности. В России статистические данные о дорожно-транспортных происшествиях и их последствиях собираются и хранятся в нескольких ведомствах: Госавтоинспекции МВД РФ, Росстата и Минздрава. Госавтоинспекция Российской Федерации ведет комплексное изучение показателей состояния безопасности дорожного движения всех областей и районов страны, в том числе Самарской области. Регулярно осуществляются измерения и обновления показателей.

В данной работе были изучены, структурированы и проанализированы материалы с различных сайтов: «Государственная статистика РФ», «Показатели состояния безопасности дорожного движения», «Федеральная служба государственной статистики по Самарской области» и других. Были отображены и сгруппированы данные, обладающие определенными признаками, а именно показатели дорожно-транспортных происшествий (ДТП) за 2015, 2016, I квартал 2017 гг., совершению которых способствовали неудовлетворительные условия содержания и обустройства улично-дорожной сети (НДУ) и пострадавшие в них люди, в том числе на пешеходных переходах, 33 районов Самарской области [1], представленные в таблице 1.

Таблица 1

Показатели ДТП, произошедших вследствие НДУ 33 районов Самарской области за 2015, 2016, I квартал 2017 гг.

№ района	Районы Самарской области	Кол-во ДТП	Кол-во пострадавших в ДТП	Кол-во пострадавших в ДТП на пешеходных переходах
1	г. Самара	1570	2052	431
2	г. Тольятти	1125	1521	317
3	г. Новокуйбышевск	85	105	30
4	г. Чапаевск	58	89	15

* © Барышева Е.Н., 2017

Барышева Евгения Николаевна (barisheva_zh@hotmail.com), кафедра математики и бизнес-информатики, Самарский национальный исследовательский университет имени академика С.П. Королева, 443086, Российская Федерация, г. Самара, Московское шоссе, 34.

Окончание табл. 1

№ района	Районы Самарской области	Кол-во ДТП	Кол-во пострадавших в ДТП	Кол-во пострадавших в ДТП на пешеходных переходах
5	Алексеевский район	4	5	0
6	Безенчукский район	17	37	1
7	Богатовский район	19	21	0
8	Большеглушицкий район	19	46	0
9	Большечерниговский район	23	34	0
10	Борский район	12	22	1
11	Волжский район	164	251	16
12	Елховский район	10	15	0
13	Исаклинский район	9	13	1
14	Камышлинский район	8	13	0
15	Кинельский район	99	163	6
16	Кинель-Черкасский район	74	134	3
17	Клявлинский район	6	8	0
18	Кошкинский район	7	7	0
19	Красноармейский район	19	40	0
20	Красноярский район	109	210	3
21	Нефтегорский район	23	34	4
22	Октябрьск	30	36	3
23	Отрадный	35	47	15
24	Пестравский район	17	23	0
25	Похвистневский район	39	56	6
26	Приволжский район	5	16	0
27	Сергиевский район	55	82	5
28	Ставропольский район	168	273	23
29	Сызранский район	292	385	88
30	Хворостянский район	11	19	0
31	Челно-Вершинский район	9	20	0
32	Шенталинский район	9	15	0
33	Шигонский район	17	22	1

С целью группировки исходного массива данных средствами кластерного анализа в работе используются показатели ДТП 33 районов Самарской области за 2015 и 2016 гг., в том числе количество пострадавших в этих ДТП в целом и конкретно на пешеходных переходах. Исходный массив данных для кластерного анализа представлен в таблице 2.

Таблица 2

Исходный массив данных

Номер района Самарской области	X_1	X_2	X_3
1	1383	1,546	0,317
2	958	1,845	0,381
3	77	0,894	0,247
4	53	1,124	0,151
5	3	0,341	0,000
6	16	0,897	0,025
7	17	1,329	0,000
8	15	1,811	0,000
9	20	1,686	0,000
10	11	0,877	0,042
11	143	2,356	0,139
12	9	1,368	0,000
13	9	1,035	0,080
14	6	1,014	0,000
15	89	4,436	0,153
16	66	2,697	0,067
17	3	0,273	0,000
18	7	0,313	0,000
19	18	2,200	0,000
20	93	3,098	0,053
21	20	0,869	0,120
22	26	1,209	0,113
23	29	0,778	0,294
24	16	1,250	0,000
25	32	1,661	0,181
26	5	0,681	0,000
27	48	1,588	0,110
28	139	2,981	0,277
29	255	13,196	2,839
30	9	0,809	0,000
31	7	0,994	0,000
32	7	0,577	0,000
33	15	0,961	0,051

X_1 – количество ДТП за 2015 и 2016 гг., совершению которых способствовали неудовлетворительные условия содержания и обустройства улично-дорожной сети;

X_2 – доля (промилле, ‰) пострадавших в ДТП вследствие НДУ, исходя из утвержденной численности постоянного населения Самарской области на 01.01.2017, утвержденной Росстатом 03.03.2017 (Приложение А).

X_3 – доля (промилле, ‰) пострадавших в ДТП вследствие НДУ на пешеходных переходах, исходя из утвержденной численности постоянного населения Самарской области на 01.01.2017, утвержденной Росстатом 03.03.2017 [2].

Используя евклидову метрику, средствами MS Excel, была рассчитана матрица расстояний размерности 33×33 . Из множества методов кластеризации в работе использовались агломеративные методы «ближнего соседа», «дальнего соседа» и «средней связи» [9]. Начало кластеризации – это поиск минимального расстояния между объектами. Из матрицы получилось, что минимальное расстояние 0,06832 – это расстояние между объектами 5 и 17. Следовательно, эти объекты образуют первый кластер $S(5;17)$. Далее осуществляется пересчет расстояния всех объектов до первого кластера $S(5;17)$ методами «ближнего соседа», «дальнего соседа» и «средней связи». Составляется новая матрица расстояний. Данный шаг повторяется до тех пор, пока матрица расстояний не станет размерности 2×2 (табл. 3, 4, 5).

Таблица 3

Окончательная таблица расстояний. Метод «ближнего соседа»

Показатели	S(1,2)	S(3:33)
S(1,2)	0	703,0959
S(3:33)	703,0959	0

Таблица 4

Окончательная таблица расстояний. Метод «дальнего соседа»

Показатели	S(1,2)	S(3:33)
S(1,2)	0	1380,001
S(3:33)	1380,001	0

Таблица 5

Окончательная таблица расстояний. Метод «средней связи»

Показатели	S(1,2)	S(3:33)
S(1,2)	0	1129,761
S(3:33)	1129,761	0

В таблице 6 отображены расстояния между объектами методами «ближнего соседа», «дальнего соседа» и «средней связи» в течение 32 этапов.

Таблица 6

Окончательная таблица расстояний

Этапы	Методы		
	«ближнего соседа»	«дальнего соседа»	«средней связи»
1	0,068	0,068	0,068
2	0,239	0,240	0,239
3	0,265	0,264	0,265
4	0,343	0,354	0,355
5	0,354	0,559	0,454
6	0,417	0,681	0,5496
7	0,826	0,826	0,826
8	0,851	0,852	0,851
9	1,0002	1,054	1,047
10	1,002	1,090	1,054
11	1,003	2,034	1,4396
12	1,054	2,057	1,562
13	1,326	2,060	2,023
14	2,002	2,405	2,235
15	2,002	3,036	3,036
16	2,029	4,051	3,144
17	2,065	4,064	3,635
18	3,036	4,219	4,051
19	3,129	5,022	4,219
20	4,001	5,089	4,573
21	4,051	6,017	4,818
22	4,219	8,023	5,022
23	5,022	11,148	10,249
24	6,0096	17,015	11,148
25	11,148	27,003	17,803
26	12,512	29,034	19,683
27	13,095	45,025	30,808
28	16,0003	90,044	57,388
29	46,001	116,477	114,517
30	112,556	252,347	153,189
31	425,0001	425,000	425,00
32	703,096	1380,001	1129,761

На рис. 1, 2, 3 изображены дендрограммы для методов «ближнего соседа», «дальнего соседа» и «средней связи» соответственно.

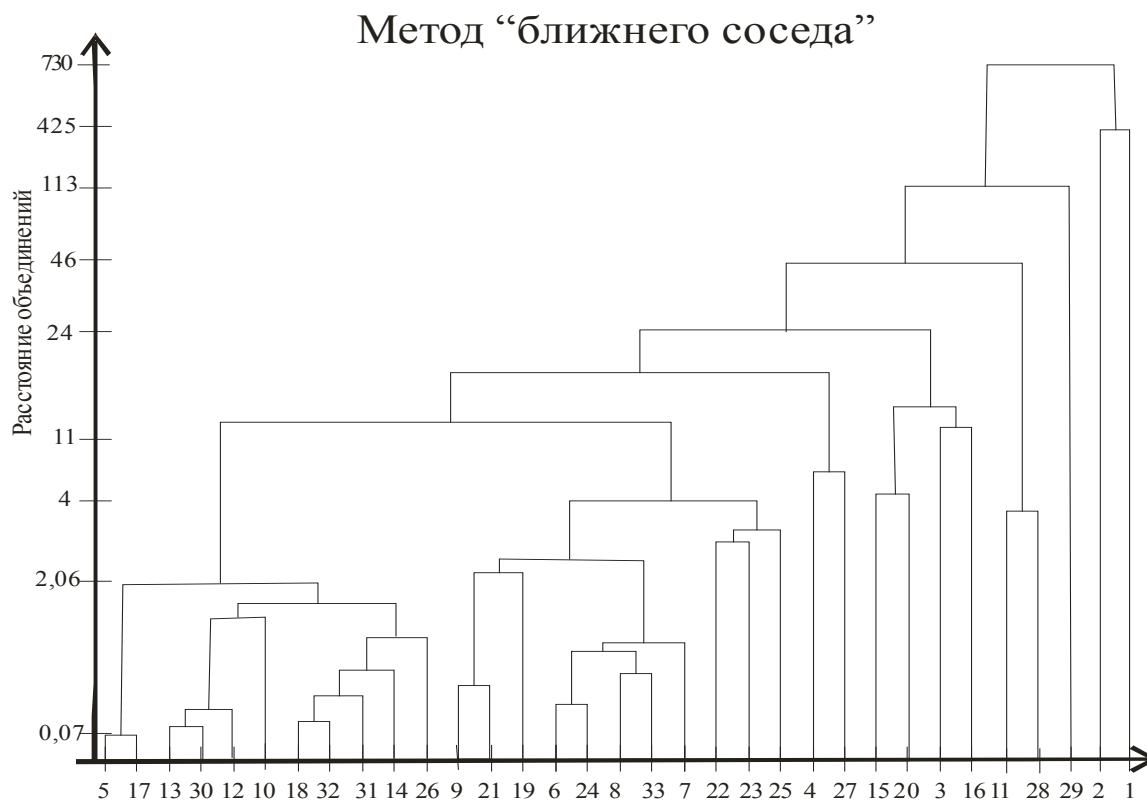


Рис. 1. Дендрограмма (метод «ближнего соседа»)

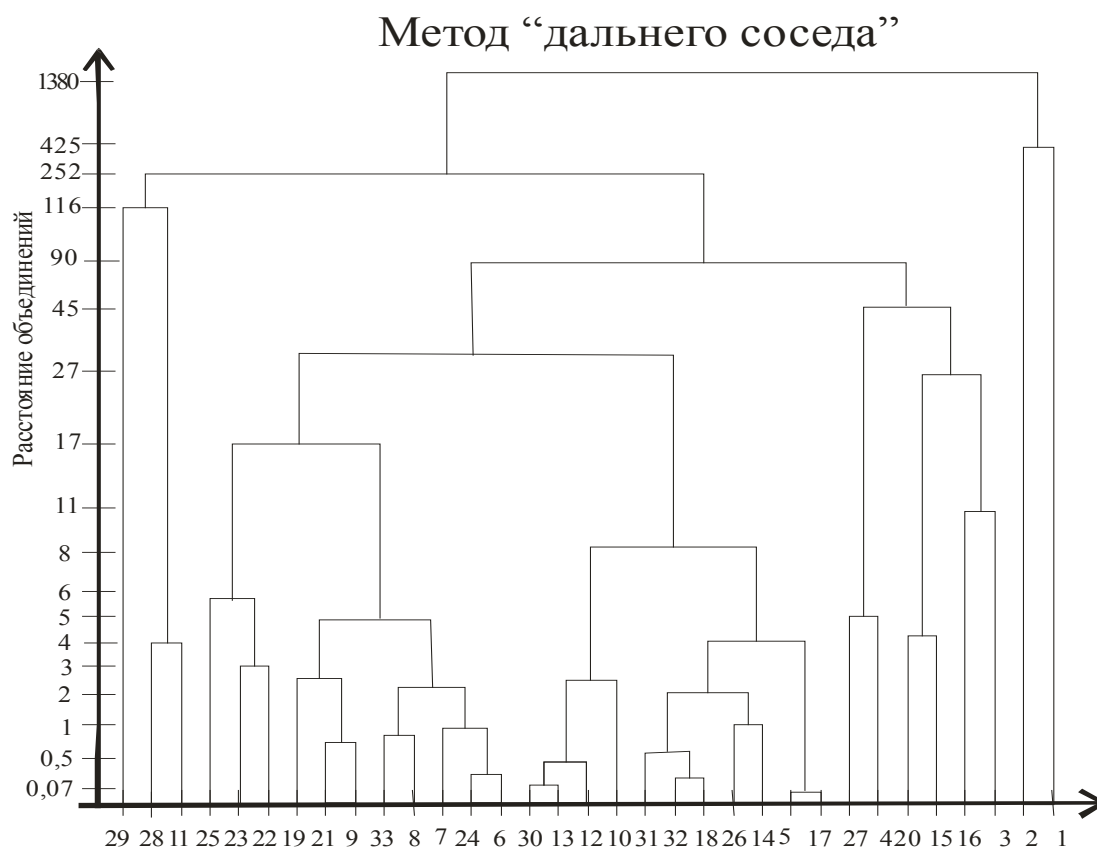


Рис. 2. Дендрограмма (метод «дальнего соседа»)

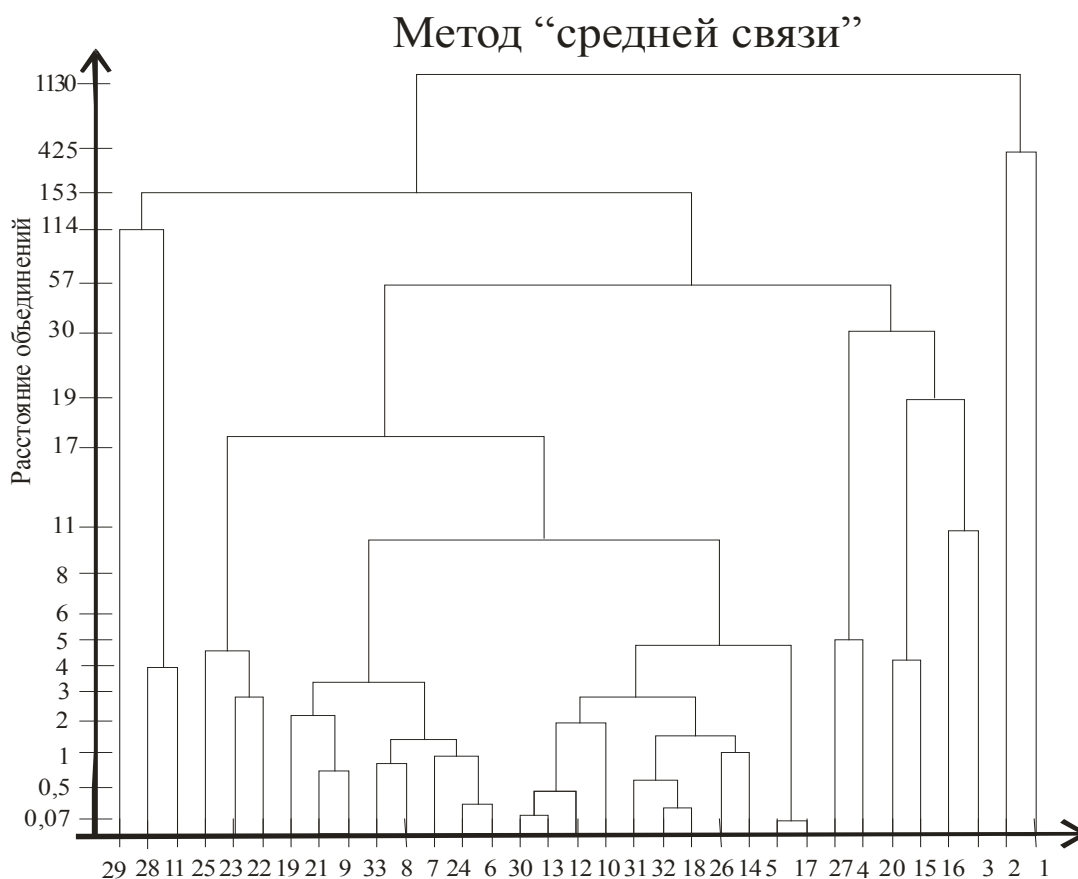


Рис. 3. Дендрограмма (метод «средней связи»)

После анализа дендрограмм, полученных в ходе классификации, были образованы кластеры: метод «ближнего соседа» – три кластера: $S(1,2,11,28,29)$, $S(3,4,15,16,20,22,23,25,27)$ и $S(5,6,7,8,9,10,12,13,14,17,18,19,21,24,26,30,31,32,33)$; метод «дальнего соседа» – четыре кластера: $S(1,2,11,28,29)$, $S(3,4,15,16,20,27)$, $S(6,7,8,9,19,21,22,23,24,25,33)$ и $S(5,10,12,13,14,17,18,26,30,31,32)$; метод «средней связи» – пять кластеров: $S(1,2,11,28,29)$, $S(3,4,15,16,20,27)$, $S(22,23,25)$, $S(6,7,8,9,19,21,24,33)$ и $S(5,10,12,13,14,17,18,26,30,31,32)$.

Для данных по 33 районам Самарской области был проведен расчет критериев качества классификации [6]. Была составлена сводная таблица для функционалов, рассчитанных для методов «ближнего соседа», «дальнего соседа» и «средней связи». Результаты представлены в таблице 7.

Таблица 7

Сводная таблица функционалов

Методы		«ближнего соседа»	«дальнего соседа»	«средней связи»
Функционалы	F1	1 284 636,605	1 280 928,840	1 280 602,643
	F2	1 284 636,605	1 280 928,840	1 280 602,643

Из сводной таблицы 9 видно, что разбиение на пять кластеров $S(1,2,11,28,29)$, $S(3,4,15,16,20,27)$, $S(22, 23, 25)$, $S(6,7,8,9,19,21,24,33)$ и $S(5,10, 12,13,14,17,18,26,30,31,32)$ является самым оптимальным, так как оба критерия классификации имеют наименьшие значения. Образовавшиеся кластеры отображены в таблице 8.

Таблица 8

Показатели ДТП 33 районов Самарской области за 2015, 2016 гг.

Номер района	Район Самарской области	Кол-во ДТП	Пострадавшие в ДТП		Пострадавшие на пешеходных переходах	
			Кол-во	Доля, ‰	Кол-во	Доля, ‰
1 кластер						
1	г. Самара	1383	1808	1,5	371	0,3
2	г. Тольятти	958	1311	1,8	271	0,4
11	Волжский район	143	220	2,4	13	0,1
28	Ставропольский район	139	215	3,0	20	0,3
29	Сызранский район	255	330	13,2	71	2,8
2 кластер						
3	г. Новокуйбышевск	77	94	0,9	26	0,2
4	г. Чапаевск	53	82	1,1	11	0,2
15	Кинельский район	89	145	4,4	5	0,2
16	Кинель-Черкасский район	66	120	2,7	3	0,1
20	Красноярский район	93	38	3,1	0	0,1
27	Сергиевский район	48	72	1,6	5	0,1
3 кластер						
22	Октябрьск	26	32	1,2	3	0,1
23	Отрадный	29	37	0,8	14	0,3
25	Похвистневский район	32	46	1,7	5	0,2
4 кластер						
6	Безенчукский район	16	36	0,9	1	0
7	Богатовский район	17	19	1,3	0	0
8	Большеглушицкий район	15	34	1,8	0	0
9	Большечерниговский район	20	30	1,7	0	0
19	Красноармейский район	18	38	2,2	0	0
21	Нефтегорский район	20	29	0,9	4	0,1
24	Пестравский район	16	21	1,2	0	0
33	Шигонский район	15	19	1,0	1	0
5 кластер						
5	Алексеевский район	3	4	0,3	0	0
10	Борский район	11	21	0,9	1	0
12	Елховский район	9	13	1,4	0	0
13	Исаклинский район	9	13	1,0	1	0
14	Камышлинский район	6	11	1,0	0	0
17	Клявлинский район	3	4	0,3	0	0
18	Кошкинский район	7	7	0,3	0	0
26	Приволжский район	5	16	0,7	0	0
30	Хворостянский район	9	13	0,8	0	0
31	Челно-Вершинский район	7	15	1,0	0	0
32	Шенталинский район	7	9	0,6	0	0

Проанализировав полученные результаты из таблицы 8, можно сделать вывод, что данные 33 районов Самарской области, разделенные на 5 кластеров, имеют свои особенности:

1) г. Самара и г. Тольятти – самые большие города в области, и, соответственно, в них за 2015 и 2016 годы произошло наибольшее количество ДТП (1383 и 958) с пострадавшими (1808 чел. и 1311 чел.), в том числе на пешеходных переходах (371 чел. и 271 чел.). Однако доля пострадавших людей, исходя из общей численности населения в этих городах, невысока и составляет 1,5 % и 0,3 % соответственно.

Максимальная доля пострадавших в ДТП (13,2 %), в том числе на пешеходных переходах (2,8 %), наблюдается в Сызранском районе.

Также в первый кластер входят Волжский и Ставропольский районы с большим количеством ДТП, произошедших в результате плохого обустройства улично-дорожной сети;

2) во второй кластер входят г. Новокуйбышевск, г. Чапаевск, Кинельский, Кинель-Черкасский, Красноярский и Сергиевский районы. Это районы с меньшим количеством ДТП по сравнению с районами первого кластера. Стоит отметить Кинельский (4,4 %), Кинель-Черкасский (2,7 %) и Красноярский районы (3,1 %), в которых доля пострадавших в ДТП достаточно велика, даже по сравнению с районами первого кластера;

3) в третий кластер входят районы, в которых количество ДТП на среднем уровне – г. Октябрьск (26), г. Отрадный (29) и Похвистневский район (32);

4) в четвертый и пятый кластеры входят районы с небольшим количеством ДТП и пострадавших в них людей. В некоторых районах пострадавших на пешеходных переходах людей нет.

Классификация районов позволяет оценить, в каких районах остро стоит вопрос о необходимости ремонта дорог и обустройства пешеходных переходов [7]. Целесообразно распределять бюджет на техническое обслуживание дорог по районам, исходя из образованных кластеров. Большие же города требуют индивидуального подхода.

По данным ГИБДД, наиболее травмоопасной категорией участников дорожного движения являются велосипедисты, поэтому Правительству Самарской области необходимо оборудовать место на проезжей части для велосипедистов, как это сделано в европейских странах [8].

Практически в каждом случае наезд на пешехода происходит из-за отсутствия разметки, дорожных знаков, ограждений на пешеходном переходе. Еще одной распространенной причиной наездов на пешеходов является неправильное расположение пешеходного перехода относительно остановки общественного транспорта. Требования к обустройству пешеходных переходов в соответствии с национальными стандартами необходимо учитывать при проектировании строительства, реконструкции, капитального ремонта и ремонта автодорог.

Библиографический список

1. Сайт «ЕМИСС Государственная статистика» [Электронный ресурс]. URL: <https://fedstat.ru> (дата обращения: 10.07.2017).
2. Сайт «Показатели состояния безопасности дорожного движения» [Электронный ресурс]. URL: <http://stat.gibdd.ru> (дата обращения: 05.07.2017).
3. Сайт «Территориальный орган Федеральной службы государственной статистики по Самарской области (сокращенное название – САМАРАСТАТ)» [Электронный ресурс]. URL: http://samarastat.gks.ru/wps/wcm/connect/crosstat_ts/samarastat/ru/statistics/population/ (дата обращения: 30.06.2017).
4. Сайт Правительства Самарской области [Электронный ресурс]. URL: <http://www.samregion.ru/economy/infrastructure/roads> (дата обращения: 03.07.2017).
5. Сайт «ГАИ» [Электронный ресурс]. URL: <http://www.lgai.ru/autonews/518132-gibdd-obnarodovala-statistiku-dtp-za-2016-god.html> (дата обращения: 10.07.2017).
6. Барышева Е.Н., Никишов В.Н. Модели оценки финансовых показателей с учетом их стохастичности и хаотичности // Вестник Самарского государственного университета. 2012. № 4 (95). С. 115–126.
7. Барышева Е.Н., Сараев А.Л., Тюкавкин Н.М. Динамическая модель запаздывания эндогенных и экзогенных инвестиций в освоение капиталовложений в производственное предприятие // Вестник Самарского государственного университета. Серия «Экономика и управление». 2015. № 9/1 (131). С. 245–251.
8. Барышева Е.Н. Исследование направлений развития конкурентоспособности аэрокосмических кластеров (статья) // Вестник Самарского государственного университета. Серия «Экономика и управление». 2015. № 9/2 (131), С. 251–261.
9. Сошникова Л.А., Тамашевич В.Н., Уебе Г., Шеффер М. Многомерный статистический анализ в экономике. М.: Юнити, 2007.

References

1. EMISS Gosudarstvennaya statistika [UIIS State statistics] [Electronic resource]. Retrieved from: <https://fedstat.ru> (accessed: 10.07.2017) [in Russian].
2. Pokazateli sostoyaniya bezopasnosti dorozhnogo dvizheniya [Indicators of the state of road safety] [Electronic resource]. Retrieved from: <http://stat.gibdd.ru> (accessed: 5.07.2017) [in Russian].
3. Territorial'nyy organ Federal'noy sluzhby gosudarstvennoy statistiki po Samarskoy oblasti (SAMARASTAT) [The territorial body of the Federal State Statistics Service for the Samara Region (SAMARASTAT)] [Electronic resource]. Retrieved from: http://samarastat.gks.ru/wps/wcm/connect/rosstat_ts/samarastat/ru/statistics/population (accessed: 30.06.2017) [in Russian].
4. Pravitel'stvo Samarskoy oblasti [Site of the Government of the Samara Region] [Electronic resource]. URL: <http://www.samregion.ru/economy/infrastructure/roads> (accessed: 03.07.2017) [in Russian].
5. GAI [Electronic resource]. Retrieved from: <http://www.lgai.ru/autonews/518132-gibdd-obnarodovala-statistiku-dtp-za-2016-god.html> (accessed: 10.07.2017) [in Russian].
6. Barysheva E.N., Nikishov V.N. Modeli otsenki finansovykh pokazateley s uchetom ikh stokhastichnosti i khaotichnosti. [Models for estimating financial indicators taking into account their stochasticity and randomness]. In: Vestnik Samarskogo gosudarstvennogo universiteta [Vestnik of Samara State University], 2012, no. 4 (95), pp.115–126 [in Russian].
7. Barysheva E.N., Saraev A.L., Tyukavkin N.M. Dinamicheskaya model' zapazdyvaniya endogennykh i ekzogennykh investitsiy v osvoyeniye kapitalovlozheniy v proizvodstvennoye predpriyatiye [Dynamic model of the delay of endogenous and exogenous investments in the development of capital investments in a production enterprise]. in: Vestnik Samarskogo gosudarstvennogo universiteta [Vestnik of Samara State University]. Series «Economics and management», 2015, no. 9/1 (131), pp. 245–251 [in Russian].
8. Barysheva E.N., Issledovaniye napravleniy razvitiya konkurentosposobnosti aerokosmicheskikh klasterov [Research of directions of development of competitiveness of aerospace clusters]. In: Vestnik Samarskogo gosudarstvennogo universiteta [Vestnik of Samara State University]. Series «Economics and management», 2015, no. 9/2 (131), pp. 251–261 [in Russian].
9. Soshnikova L.A., Tamashevich V.N. Mnogomernyy statisticheskiy analiz v ekonomike [Multidimensional statistical analysis in the economy]. M.: YUNITI, 2007 [in Russian].

*E.N. Barysheva**

THE STUDY OF DEPENDENCE OF INDICATORS CHARACTERIZING THE ROAD TRANSPORT ACCIDENTS OF AREAS OF THE SAMARA REGION BY MEANS OF CLUSTER ANALYSIS

The article gives an assessment of road accidents, the fulfillment of which contributed to the unsatisfactory conditions of maintenance and arrangement of the street-road network in 33 districts of the Samara region of cluster analysis methods.

Key words: accident, Samara oblast, regions, cluster analysis.

Статья поступила в редакцию 20/VIII/2017.
The article received 20/VIII/2017.

* *Barysheva Evgeniya Nikolayevna* (barisheva_zh@hotmail.com), Department of Mathematics and Business Informatics, Samara National Research University, 34, Moskovskoye shosse, Samara, 443086, Russian Federation.